Check for updates

Original Research

SAGE Open
January-March 2024: 1–23
© The Author(s) 2024
DOI: 10.1177/21582440231211802
journals.sagepub.com/home/sgo
Sage

A Partial Least Squares Structural Equation Modeling Exploration of EFL Learners' Perceptions of What Contributes to the Readability of Model Paragraphs

Tuyen Thanh Nguyen[1], John R. Baker[2], and Thao Quang Le[3]

Abstract

This study investigated what features undergraduate EFL learners perceive as affecting the difficulty of model paragraphs. Four hundred and seventy-five Vietnamese undergraduates participated in a partial least squares structural equation model design. They ranked five paragraphs from easiest to most difficult and responded to a 10-point Likert questionnaire regarding 11 features (titles, paragraph length, vocabulary, vocabulary in context, rhetorical organization, paragraph structure, sentence length, punctuation, signal words, interest, background knowledge). The results showed that eight variables (titles, vocabulary, vocabulary in context, sentence length, rhetorical organization, paragraph structure, interest, background knowledge) had a significant direct effect and four variables (vocabulary, sentence length, rhetorical organization, background knowledge) had mediating effects. The model accounted for 0.508 R^2 of students' perceptions, with a moderate to high predictive relevance ($Q^2 = 0.35$). The paper also discusses the results' implications for those in writing studies and the publishing industry. Suggestions for future study are also presented.

Keywords

model paragraphs, readability, partial least squares structural equation modeling, PLS-SEM, checklist, textbooks, English as a foreign language

Introduction

Publishers include model paragraphs in textbooks because genre-specific reading has been shown to facilitate students' writing (Hyland, 2007) in what has been termed the reading-writing relationship (Shanahan & Lomax, 1988): Reading facilitates better writing (Thaiss & Zawacki, 2006). However, students' opportunity to garner these benefits is hampered if they cannot grasp what they are reading. In short, students cannot learn from what they cannot read (Allington, 2002). As such, educators must consider whether the materials are a good match for intended readers (Baker, 2019), the study of which is known as readability (Gilliland, 1972).

Readability assessment has been widely researched and applied in the past century (DuBay, 2007a) by applying two-factor (semantic, syntactic) quantitative readability formulae to measure texts, as these two features have been shown to be reliable predictors of readability and easily measured (DuBay, 2007b). However, examining only two factors and how they apply to the text, not exploring reader-text interaction, has been repeatedly criticized as overly reductive, as it has become generally accepted that readability assessments should include a consideration of readers' perceptions of the many features that make up text difficulty (Baker, 2021; Gunning, 2003; Weaver, 2000).

In response, comprehensive lists of features readers perceive as affecting the difficulty of various texts have

[1]EQuest Education Group, Ho Chi Minh City, Vietnam
[2]Ton Duc Thang University, Ho Chi Minh City, Vietnam
[3]Van Lang University, Ho Chi Minh City, Vietnam

Corresponding Author:
John R. Baker, Creative Language Center, Ton Duc Thang University, 19 Nguyen Huu Tho Street, Tan Phong Ward, Ho Chi Minh City, District 7, Vietnam.
Email: drjohnrbaker@tdtu.edu.vn

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

been offered for use in what has been termed a hybrid procedure: Employing a readability formula is considered a good first step, followed by a second step that includes a subjective consideration of features not measured by readability formulae as this is intended to provide texts that are a good fit for potential readers (Chall & Dale, 1995; Fry, 2002; Gunning, 2003; Meyer, 2003; Weaver, 2000).

Some of these lists have been developed for general texts and native English speakers (NES) (Chall & Dale, 1995; Zakaluk, 1985; Zakaluk & Samuels, 1988). Others have been designed for more genre-specific texts (i.e., model essays) and English as a foreign language (EFL) learners (Baker, 2020). A common thread among these lists is that they are often inspired by previous research that has identified one or more features that contribute to readers' difficulty with texts.

Literature Review

Second-language (L2) research often finds girding in first-language (L1) scholarly precedents. Aligned with this, L1 discussions began in the early 1900s, and L2 literature followed in the 1970s and 1980s. These theoretical and empirical explorations focused on a sole or a limited number of primary features and a small number of conjoined relationships that affect readability. In keeping with this trajectory, the literature is presented as such.

Titles

Theoretical literature on titles (a descriptor at the top of a text) usually references Bartlett's (1932) work with schema. Following this, landmark empirical explorations in NES contexts showed titles are facilitative as they forecast the topic of a text (Bransford & Johnson, 1973; Dooling & Lachman, 1971). Research that followed in EFL contexts has reported similar results (Carrell, 1983; Noor, 2006), as readers often first look to titles when approaching a text. The title has also been found to mediate the effect of other features, helping readers to increase interest (Mohammed, 2021), activate background knowledge (Ahmadi, 2011), and anticipate text structure (Bock, 1980) and rhetorical organization (Baker, 2020).

Paragraph Length

Discussions of paragraph length (the number of words in a text) began in the 1890s (Earle, 1890; Lewis, 1894) and have historically shown that longer narrative texts are better comprehended and recalled than shorter ones as additional details (subsidiary sentences) strengthen plots (Keenan et al., 1985; Mandler & Johnson, 1977). Expository explorations, however, have been less conclusive. Some suggest additional details overburden readers' memory (Reder, 1982; Reder & Anderson, 1980). Others have reported the opposite (Reder et al., 1986). Still, others have shown that readers overestimate their comprehension of shorter texts (Commander & Stanwyck, 1997).

EFL explorations have similarly been contradictory. Several have shown a relationship between narrative text difficulty and length (Gopal & Mahmud, 2019), but others failed to establish a relationship (Jalilehvand, 2012). Similarly, several expository studies have found that text length increases difficulty (Freedle & Kostin, 1991, 1992, 1993; Moon, 2019), while others reported no effect (Lee, 1999; Mehrpour & Riazi, 2004). Mediating effects have also been observed in EFL research. Longer texts, for example, have been shown to contain more vocabulary (Hung, 2017), thus negatively affecting comprehension (Bock, 1980) and interest (Baker, 2020). Moreover, vocabulary in context clues have been reported to be helpful in shorter texts (Shokouhi & Askari, 2010). However, shorter passages provide less context, requiring more reliance on background knowledge. Similarly, reduced text structures require increased reader effort to understand the text (Bae & Lee, 2018).

Vocabulary

Vocabulary (i.e., unfamiliar, abstract, figurative, or technical words) has regularly been cited as a feature that affects readers' understanding since the early 1920s (Pressey & Pressey, 1921) and a contributing factor to efficient, silent reading and other reading skills (Davis, 1944).

Studies with EFL learners have similarly demonstrated that vocabulary plays a role in text difficulty, showing that EFL learners perceive it to be their greatest obstacle to text comprehension and recall (Kameli & Baki, 2013; Kezhen, 2015; Qian, 2002; Salyer, 1990; Yorio, 1971). Vocabulary has also been found to interact with other features. That is, difficult vocabulary can reduce the assistance vocabulary in context clues provide, be more abundant in longer sentences (Haynes & Baker, 1993), reduce interest (Baker, 2021), and hinder recognition of rhetorical organization (Carrell, 1983).

Vocabulary in Context

Discussions of vocabulary in context (how effectively phrases or sentences that surround unknown words aid comprehension) in NES contexts often begin with Ames's (1966) categorization of textual clues and how readers use these to infer meaning of unfamiliar words, which in turn facilitates comprehension.

Explorations in EFL contexts have also resulted in classification systems (Bengeleil & Paribakht, 2004; Dubin & Olshtain, 1993) and shown that EFL learners use contextual clues to guess the meaning of unknown words (Ahmad et al., 2018; Cooper, 1999). Some clues, however, have proven more assistive than others. Immediate clues, rather than global ones, have been shown to be more conducive to learning (Haynes, 1993), and clues with limited redundancy and limited or ambiguous references have been shown to be less so (Laufer, 1997). Additionally, there is evidence that learners' abilities to recognize clues play a part (Shen & Wu, 2009) and improve with training (Davoudi & Nafchi, 2016; Rokni & Niknaqsh, 2013). Vocabulary in context has also been shown to have a relationship with other features but one that is affected by them (see vocabulary, background knowledge, and paragraph length sections of this article).

Sentence Length

Discussions of sentence length (the number of words in a sentence) often begin with Sherman (1893), who stressed that readable materials do "not run in long and involved sentences that cannot readily be understood" (p. 327) and show that longer sentences tend to be more difficult as they often are more complex (including longer clauses, more abstract nouns, verb nominalizations, adjectives, dependent clauses, and adverbials) and the accompanying punctuation (Coleman, 1962; Coleman & Miller, 1968; Glazer, 1974), which can overtax readers' working memories (McLaughlin, 1969; Mikk, 2008) or generally result in a misunderstanding (McElree, 2000) (e.g., overtax the cognitive load needed for the recognition of paragraph structure and rhetorical organization).

EFL literature has provided similar results (Freedle & Kostin, 1992) but has added that readers' proficiency can play an important role (Nilagupta, 1977), as much as a .64 correlation (Dwaik, 1997). Sentence length has also been shown to interact with other features. For example, shorter sentences can increase interest (Mikk & Kukemelk, 2010).

Rhetorical Organization

Discussions of rhetorical organization (the way ideas are organized in texts to make them flow smoothly) generally begin with Meyer's (1975) subclassifications and include more modern typologies (e.g., illustration, process, description, narrative, cause/effect, comparison/contrast, and argumentation/persuasion) (Baker, 2021), but one type can occur in another (Spiro & Taylor, 1980). In general, two sources of difficulty have been noted: (a) the complexity of rhetorical classification/subclassification and (b) readers' formal schemata, and familiarity with the rhetorical organization employed (Carrell, 1987).

Similar discussions have been offered in EFL contexts. That is, some text types are more complex than others, learners' awareness can affect understanding (Flick & Anderson, 1980), and rhetorical structures are not constant across cultures (Kaplan, 1966, 2005). Nevertheless, taxonomy explorations have attempted to provide as distinctive a picture as possible (see Alkhaleefah, 2017; Amiri et al., 2012; Baker, 2021; Carrell, 1984a; Freedle & Kostin, 1991, 1993; Goh, 1990; Lei, 2010; Meyer & Freedle, 1984; Putra, 2012; Saadatnia et al., 2016; Salmani, 2010; Sharp, 2002; Talbot et al., 1991; Yali & Jiliang, 2007; Zhang, 2008). However, the results have generally been incongruent due mostly to varying text types under study, methodology, and participants' reading levels (Baker, 2021). Rhetorical organization has also been shown to interact with other features. That is, it influences students' use of vocabulary in context clues (Baker, 2020).

Structure

Discussions of text structure (how text is organized) began in the 1970s, showing that narrative texts with a well-defined story grammar facilitate understanding better than those without (Thorndyke, 1975). Likewise, readers expect a clear structure for expository texts. For example, a text with an identifiable topic sentence (Lorch & Lorch, 1985) and one that follows a conventional development structure aligned with the relevant rhetorical style (Britton & Black, 2017; Kintsch & Yarbrough, 1982). And when this is not met, comprehension suffers (Kieras, 1978).

Similar results have been found in EFL contexts. It has been shown that text structure can influence readers' experiences (Baker, 2020) as learners come with a predisposed schema regarding narrative (Carrell 1984b) and expository structures (Ritzer, 1994). Thus, how well a text's structure meets these expectations can affect comprehension (Carrell, 1992). This relationship can also be influenced by readers' proficiency (Walters & Wolf, 1986) and awareness and knowledge of conventional text structures (Namjoo & Marzban, 2012; Shemshadsara et al., 2019).

Structure has also been shown to have a relationship with other features, but one where structure is mediated by them (see the title section of this article).

Signal Words

Discussions of signal words (words that indicate the flow of information, e.g., first, next, finally, etc.) often begin

with Thorndike (1917), who demonstrated that signal words can facilitate NES comprehension, and continue with Miccinati (1975), who organized them into several categories that have become integral to writing studies courses (Baker, 2021; Van Silfhout et al., 2014). Evidence regarding signal words' effects on comprehension have been mixed. Some early explorations have reported that signal words support reading comprehension (Miccinati, 1975), whereas others have demonstrated negative effects (Roen, 1984), and still others have indicated no effect (Meyer, 1975).

However, research with EFL learners begins with Le (1969) and has shown mostly positive results (i.e., signal words facilitate understanding) (Aidinlou & Pandian, 2011; Al-Surmi, 2011), as they support a coherent mental representation of clause relations (Xu et al., 2019). However, learner proficiency (Chung, 2000; Kim & Clariana, 2017) and awareness (Baker, 2021; Quan, 2008) have been shown to play a part. Signal words have also been shown to interact with other factors. That is, signal words can aid recall and recognition of rhetorical organization and structure (Baker, 2022; Lorch & Chen, 1986).

Punctuation

Discussions of punctuation (periods, question marks, exclamation marks, commas, colons, semicolons, dashes/hyphens, ellipsis, etc.) begin with Summey's (1919) handbook, which detailed modern uses of punctuation and how it aids comprehension (Backscheider, 1972). However, the degree of assistance provided is a balance of the presence of punctuation (Neff, 1932) and readers' understanding of its purpose and standard usage (Carr, 1978; Carver, 1970; Durkee, 1952), without which sentences can be no more than a jumble of text (Hasbrouck et al., 1999).

Similar arguments have been made in EFL contexts, explaining that punctuation is facilitative, but student awareness plays an important part (Abbott, 2006; Alsubaie, 2014; Benitez-Rivera, 2013; Pathan & Al-Dersi, 2013; Shih, 1992; Suliman et al., 2019). Punctuation has also been shown to have a relationship with other features, but one where it is mediated by other components (see sentence length and background knowledge).

Interest

Discussions of interest (an interest in a topic) begin with Hebart's work from the 1800s (Dewey, 1913), who explained that interest or absence thereof affects learning. Later work reified this point (Lin et al., 1997; Schraw & Lehman, 2001).

Research in EFL contexts has similarly demonstrated that students who express topic interest (Atamturk, 2018) and those who do not may be interested in learning more (Baker, 2020; Erçetin, 2010). Interest has also been shown to have a relationship with other features, but one where it is mediated by other components (see titles, vocabulary, sentence length, paragraph length, and background knowledge sections of this article).

Background Knowledge

Discussions of background knowledge (how familiar students are with the topic) often begin with Kant's 1781 treatise on schemata (Scaglia, 2020) and Bartlett's (1932) schema classifications. EFL explorations additionally reference Carrell's (1983) content schema, which posits that a text does not provide meaning but only provides guidance filtered by readers' previous background knowledge (Carrell, 1983, 1987; Chau et al., 2019; Florencio, 2004; Ghorbandordinejad & Bayat, 2014; Ha, 2021; Khataee & Davoudi, 2018; Nelson, 1987; Nguyen, 2012; Steffensen et al., 1979; Thao & Son, 2018).

Background knowledge has also been shown to interact with other features: vocabulary (Johnson, 1981; Sheridan et al., 2019); vocabulary in context (Demir, 2012; Johnson, 1982), sentence length and punctuation (Johnson, 1981), rhetorical organization, structure (Carrell, 1987), and interest (Ay & Bartan, 2012; Bugel & Buunk, 1996; Carrell & Wise, 1998; Kelsen, 2016).

Aim of the Study

A review of the extant literature illustrates that explorations of one or more features' effects on passage difficulty have been undertaken. However, limitations are similarly present, as these investigations have only explored one or a small number of variables and a limited number of mediating relationships. Additionally, research into what primary and mediating features EFL learners perceive as contributing to the difficulty of model paragraphs is noticeably absent. This study is intended to address this combined gap. Pursuant to this aim, two research questions were posed:

> RQ1: What factors do undergraduate EFL learners perceive as affecting the text difficulty of model paragraphs?
>
> RQ2: What mediating factor relationships do undergraduate EFL learners perceive as affecting the text difficulty of model paragraphs?

To explicate these two questions, 11 hypotheses (and relevant sub hypotheses) were posed using a transmittal approach, testing path relationships based on existing literature (Nitzl et al., 2022). RQ1 is explicated by the main

hypotheses (H1-H11), and RQ2 is explicated by the sub-hypotheses.

H_1 There is a significant relationship between titles and participants' perceptions of text difficulty.
H_{1a} Titles mediate the relationship between rhetorical organization and participants' perceptions of text difficulty.
H_{1b} Titles mediate the relationship between paragraph structure and participants' perceptions of text difficulty.
H_{1c} Titles mediate the relationship between background knowledge and participants' perceptions of text difficulty.
H_{1d} Titles mediate the relationship between interest and participants' perceptions of text difficulty.
H_2 There is a significant relationship between paragraph length and participants' perceptions of text difficulty.
H_{2a} Paragraph length mediates the relationship between vocabulary and participants' perceptions of text difficulty.
H_{2b} Paragraph length mediates the relationship between vocabulary in context and participants' perceptions of text difficulty.
H_{2c} Paragraph length mediates the relationship between paragraph structure and participants' perceptions of text difficulty.
H_{2d} Paragraph length mediates the relationship between interest and participants' perceptions of text difficulty.
H_{2e} Paragraph length mediates the relationship between background knowledge and participants' perceptions of text difficulty.
H_3 There is a significant relationship between vocabulary and participants' perceptions of text difficulty.
H_{3a} Vocabulary mediates the relationship between vocabulary in context and participants' perceptions of text difficulty.
H_{3b}Vocabulary mediates the relationship between sentence length and participants' perceptions of text difficulty.
H_{3c} Vocabulary mediates the relationship between rhetorical organization and participants' perceptions of text difficulty.
H_{3d} Vocabulary mediates the relationship between interest and participants' perceptions of text difficulty.
H_4 There is a significant relationship between vocabulary in context and participants' perceptions of text difficulty.
H_5 There is a significant relationship between sentence length and participants' perceptions of text difficulty.
H_{5a} Sentence length mediates the relationship between rhetorical organization and participants' perceptions of text difficulty.
H_{5b} Sentence length mediates the relationship between paragraph structure and participants' perceptions of text difficulty.
H_{5c} Sentence length mediates the relationship between punctuation and participants' perceptions of text difficulty.
H_{5d} Sentence length mediates the relationship between interest and participants' perceptions of text difficulty.
H_6 There is a significant relationship between rhetorical organization and participants' perceptions of text difficulty.
H_{6a} Rhetorical organization mediates the relationship between vocabulary in context and students' perceptions of text difficulty.
H_7 There is a significant relationship between paragraph structure and participants' perceptions of text difficulty.
H_8 There is a significant relationship between signal words and participants' perceptions of text difficulty.
H_{8a} Signal words mediate the relationship between rhetorical organization and participants' perceptions of text difficulty.
H_{8b} Signal words mediate the relationship between paragraph structure and participants' perceptions of text difficulty.
H_9 There is a significant relationship between punctuation and participants' perceptions of text difficulty.
H_{10} There is a significant relationship between interest and participants' perceptions of text difficulty.
H_{11} There is a significant relationship between background knowledge and participants' perceptions of text difficulty.
H_{11a} Background knowledge mediates the relationship between vocabulary and participants' perceptions of text difficulty.
H_{11b} Background knowledge mediates the relationship between vocabulary in context and participants' perceptions of text difficulty.
H_{11c} Background knowledge mediates the relationship between rhetorical organization and participants' perceptions of text difficulty.
H_{11d} Background knowledge mediates the relationship between paragraph structure and participants' perceptions of text difficulty.
H_{11e} Background knowledge mediates the relationship between punctuation and participants' perceptions of text difficulty.
H_{11f} Background knowledge mediates the relationship between interest and participants' perceptions of text difficulty.

To explore the hypotheses and sub hypotheses, a partial-least squares structural equation (PLS-SEM) model was posed and tested. A PLS-SEM design was employed to address the limitations of previous studies (e.g., investigating a limited number of features and relationships), as sole hypotheses are individual conjectures, whereas PLS-SEM enables the investigation of complex models with many constructs and mediating relationships in a causal predictive approach that emphasizes prediction in estimating statistical models where the structure is designed to provide causal explanations of the relationships among constructs (Hair et al., 2018; Wong, 2019).

Methods

Drawing on the extant literature, a PLS-SEM model was developed to explore the two research questions and related hypotheses. The model contained 11 unobservable independent variables (constructs): Participants' perceptions regarding how 11 features affected their perceptions of the dependent variable text difficulty (TD): titles (T), paragraph length (PL), vocabulary (V), vocabulary in context (VC), sentence length (SL), rhetorical organization (RO), paragraph structure (PS), signal words (SW), punctuation (P), interest, (I), and background knowledge (BK). Drawing further on the literature and resulting hypotheses, indirect mediating relationships were additionally explored.

Setting and Participants

The study was conducted at Van Lang University in Ho Chi Minh City, Vietnam. A nonprobability method was employed to identify the target sample: the entire cohort of Writing II (17 sections) who had completed Writing 1 and can be expected to be familiar with the type of paragraph genre explored in this study, 734 potential participants. Due to Covid-19 infections, 259 students were absent. No follow-up was attempted. Four hundred and seventy-five surveys were collected.

Materials

The paragraphs under study were excerpted from the first-year composition course reference text (Savage & Shafiei's, 2012 Effective Academic Writing 1: The Paragraph, 2nd ed). The text contains 12 model paragraphs, of which five were purposefully chosen. The number is large enough to provide a wide variety of comparative options for participants to make insightful comparisons but small enough for participants to scrutinize them in a reasonable amount of time to collect meaningful data, without undue participant fatigue (Baker, 2020).

The paragraphs were chosen to be near in difficulty (as measured by the DRP formula) so as not to make the ranking obvious. Each paragraph contained the aforementioned characteristics in varying degrees but was not specifically selected as such to avoid influencing the results (Table 1) (Baker, 2020).

Experimental Procedures

After the paragraphs' identification, a cline-questionnaire was administered (O'Hear et al., 1992). In the cline phase, participants ranked the paragraphs from easiest to most difficult. To facilitate the sort of decision-making process usually used to make such judgments, the paragraphs were provided in random order without ranking criteria (Chall et al., 1996). The Friedman test was used to determine the means and significance of participants' rankings.

In the following phase, participants completed a 10-point Likert questionnaire to provide insight into their perceptions of factors contributing to text difficulty. The survey was developed from existing literature (Dörnyei & Taguchi, 2009) and reviewed by field experts ($N = 3$) to ensure clarity and relevance to each construct, and a PLS-SEM expert was consulted. To ensure reliability, the questionnaire was translated into the students' native language (Vietnamese) via back-translation and checked by a second translator. The translation and administration were then evaluated by fully bilingual TESOL instructors ($N = 3$) with over 3 years' experience and piloted with a small number of students ($N = 10$).

To attain a high response rate, the cline-questionnaire procedure was conducted during regular class periods (Brown, 2001; Kropf & Blair, 2005). To motivate students to participate, reduce nonresponse bias and missing data, and improve overall response quality, several motivators were addressed (e.g., altruistic motivation and interest) (Singer & Ye, 2013). A token incentive was also provided in the event other motivators were not present (10,000 Vietnam Dong phone card, approximately 40 US cents): Small enough to compensate for time spent and inconvenience while not being unethically compelling (Ripley et al., 2010).

Prior to the study, an exploratory factor analysis (EFA) was conducted with 174 undergraduates with similar demographics and experience with the text. The results showed that each variable loaded well on its factor and that the results were significant (Bartlett's Test of Sphericity, $p = .000$) (Table 2) and had high sampling adequacy (Kaiser-Meyer Olkin Measure, 0.795) (Table 3).

The Cronbach alpha (Cα) for each factor was found to be above .70. This indicated acceptable scale reliability (Table 4).

Table 1. Features of the Texts.

	The long life of my grandfather's car	Something wild	My brother's game	St. Petersburg	The secret to a successful vacation
DRP	49	52	53	54	55
Title	Yes	Yes	Yes	Yes	Yes
Paragraph length	179 words	261 words	158 words	203 words	227 words
Vocabulary*	1.0%	.66%	1.64%	4.4	1.72%
Vocabulary in context	Some clues may be unhelpful	Some words need to be inferred from the surrounding context	Some words need to be inferred from the surrounding context	May have trouble inferring the meaning of unfamiliar words from the surrounding context	Inferring vocabulary from the context is not necessary
Sentence length	11.9 words	14.5 words	10.5 words	12.6 words	11.9 words
Rhetorical organization	Descriptive	Narrative	Example	Opinion	Process
Paragraph structure	Well organized	Moderately well organized	Pretty well organized	Well organized	Well organized
Signal words	Yes, e.g., because, when, but, after	Yes, e.g., for my 25th birthday, because, on the day of, etc.	Yes, e.g., but, because, one of, so, the other, etc.	Yes, e.g., first, in addition, the third, finally, etc.	Yes, e.g., the first step is, next, now, such as, etc.
Punctuation	Yes, e.g., 15 periods, 5 commas, and 2 apostrophes	Yes, e.g., 13 commas, 17 periods, 1 question mark, 2 ellipsis marks, 1 quotation mark set, and 1 hyphen	Yes, e.g., 15 periods and 5 commas	Yes, e.g., 18 periods, 18 commas, and 1 apostrophe	Yes. e.g., 17 periods, 2 exclamation marks, and 8 commas
Interest	Yes*	Yes	Yes	Yes	Yes
Background knowledge	Yes	Yes	Yes	Yes	Yes

*Found on Coxhead's (2000) Academic Word list (offered as reference).

Table 2. Kaiser-Meyer-Olkin Measure.

Kaiser-Meyer-Olkin Measure of sampling adequacy	0.795

Table 3. Bartlett's Test of Sphericity.

Approx. Chi-Square	*df*	Sig.
11,727.144	2556	.000

Table 4. Cronbach's alpha.

Construct	Cα	Construct	Cα	Construct	Cα
BK	.925	PS	.894	T	.859
I	.927	RO	.914	TD	.925
PL	.851	SL	.899	V	.861
P	.926	SW	.906	VC	.881

After the completion of the pilot study, the study was undertaken. Following the data collection, two research assistants independently entered the survey responses into Microsoft Excel CSV files, which were then verified by a third. Several data preparation issues were addressed. (a) missing data (MCAR, expectation-maximization algorithm) (Haziza, 2009), (b) suspicious response patterns (straight-lining or inconsistent answers), (c) outliers, and (d) abnormal data distribution (Hair et al., 2021). Four hundred and forty-one usable responses were set for analysis, a sample size larger than PLS-SEM suggested specifications. That is, the priori sample method (Soper, 2022) and the 10 Item Method (Hair et al., 2017). No participants were excluded based on demographic characteristics. These included participants of varying gender, ages, and year of study (Table 5).

The PLS-SEM outer measurement model was then examined using Indicator Reliability, Internal Consistency, and Discriminant Validity (Fornell-Larcker Criterion, Cross Loading, and Heterotrait-Monotrait Ratio—HTMT). The inner structure model was assessed using Collinearity, Path Coefficients, Mediating Relationships, Explanatory Power (Coefficients of Determination, R^2), and Predictive Relevance (Q^2).

Results

The Friedman Test (mean rank table) (Table 6) demonstrated that My Brother's Game paragraph was reported to be perceived to have the lowest average difficulty, followed by The Long Life of My Grandfather's Car, The Secret to a Successful Vacation, St. Petersburg, and Something Wild.

Table 5. Demographics.

	N	(%)
Gender		
Male	150	34.0
Female	281	63.7
Other	4	0.9
Prefer not to say	6	1.4
Age group		
18	204	46.3
19	181	41.0
20	20	4.5
21	23	5.2
22	7	1.6
23	4	0.9
24	0	0.0
25	2	0.5
Year of study		
1	420	95.2
2	6	1.4
3	12	2.7
4	3	0.7

Table 6. Paragraph Cline (Rankings).

Paragraphs	Mean	*SD*	Min	Max
My brother's game	1.42	0.80	1.00	5.00
The long life of my grandfather's car	2.72	1.08	1.00	5.00
The secret to a successful vacation	3.31	1.31	1.00	5.00
St. Petersburg	3.58	1.19	1.00	5.00
Something wild	3.97	1.07	1.00	5.00

Table 7. Friedman Test.

Chi-Square	*df*	Asymp. sig.
694.70	4	.000

A significant difference in the ranking of each essay was shown ($\chi 2(4)$, $p < .001$), thus demonstrating that the informants made definitive choices in their rankings (Table 7).

The assessment of the outer measurement model included several areas: Indicator Reliability, Internal Consistency, and Discriminant Validity (Fornell-Larcker Criterion, Cross Loading, and HTMT).

An examination of the indicator loadings showed 66 of the 72 indicators were above the 0.70 threshold (Table 8)

Table 8. Indicator Reliability.

Construct	Items	Standardized loadings	Construct	Items	Standardized loadings	Construct	Items	Standardized Loadings
BK	BK1	0.739	PS	PS1	0.76	T	T1	0.744
	BK2	0.882		PS2	0.87		T2	0.829
	BK3	0.838		PS3	0.863		T3	0.811
	BK4	0.786		PS4	0.757		T4	*
	BK5	0.877		PS5	0.756		T5	*
	BK6	0.79		PS6	0.774		T6	0.707
I	I1	0.732	RO	RO1	0.779	TD	TD1	0.804
	I2	0.82		RO2	0.875		TD2	0.848
	I3	0.822		RO3	0.856		TD3	0.867
	I4	0.783		RO4	0.799		TD4	0.748
	I5	0.824		RO5	0.795		TD5	0.83
	I6	0.751		RO6	0.815		TD6	0.738
P	P1	0.763	SL	SL1	0.819	VC	V1	0.733
	P2	0.887		SL2	0.886		V2	0.873
	P3	0.873		SL3	0.896		V3	0.894
	P4	0.826		SL4	0.839		V4	*
	P5	0.799		SL5	0.742		V5	*
	P6	0.845		SL6	0.83		V6	0.753
PL	PL1	0.767	SW	SW1	0.772	VC	VC1	0.785
	PL2	0.861		SW2	0.837		VC2	0.83
	PL3	0.852		SW3	0.836		VC3	0.866
	PL4	0.726		SW4	0.789		VC4	0.705
	PL5	*		SW5	0.791		VC5	*
	PL6	0.767		SW6	0.799		VC6	0.735

*Removed from the measurement scale because they were below the threshold of 0.70.

(i.e., six were not). As such, those below the threshold were removed from the measurement scale: PL5 (0.698), T4 (0.626), T5 (0.653), V4 (0.668), V5 (0.657), and VC5 (0.635).

Three methods were used to establish reliability. Cronbach's alpha and Composite Reliability (CR) (Hair et al., 2017) showed all constructs were above the required 0.70 threshold (Hair et al., 2011). The third indicator, rhoA, showed that all loadings were between 0.70 or higher but lower than 1 (Wong, 2019). Finally, convergent validity was established at the recommended Average Variance Extracted Value (AVE) greater than or equal to 0.50, indicating that items converged to measure the underlying construct (Table 9).

Three methods were used to evaluate discriminant validity: Fornell and Larcker (1981) Criterion, cross-loadings, and HTMT. Fornell-Larcker demonstrated discriminant validity, as the square root of the AVE for each construct was greater than its correlation with all other constructs (Table 10).

Cross-loadings indicated that each indicator loaded strongly onto its parent construct and not on other constructs (Hair et al., 2017). This further demonstrated discriminant validity (Table 11).

The HTMT ratio was assessed at 0.85 or less (Kline, 2011). Results were below 0.85 (Table 12), further demonstrating construct validity.

Table 9. Internal Consistency.

Construct	Cα	CR	rhoA	AVE
BK	.902	0.925	0.908	0.673
I	.879	0.908	0.880	0.623
P	.911	0.931	0.915	0.694
PL	.854	0.896	0.853	0.633
PS	.885	0.913	0.890	0.637
RO	.903	0.925	0.905	0.673
SL	.913	0.933	0.914	0.700
SW	.891	0.916	0.893	0.647
T	.777	0.857	0.783	0.600
TD	.892	0.918	0.896	0.652
V	.829	0.888	0.834	0.666
VC	.844	0.890	0.848	0.618

The inner structural model was evaluated through an examination of several areas: collinearity, significance of the structural model relationship path coefficients, mediation analysis, explanatory power: coefficients of determination (R^2), and predictive relevance (Q^2).

An examination of collinearity showed a variance inflation factor (VIF) below 5.0 for all constructs (Table 13). As such, no sign of excessive collinearity was found among the predictor constructs (Hair et al., 2021) (Table 13).

Afterward, 11 hypotheses (H_1-H_{11}) were examined. These hypotheses queried the direct relationships

Table 10. Discriminant Validity—Fornell-Larcker Criterion.

	BK	I	P	PL	PS	RO	SL	SW	T	TD	V	VC
BK	***0.820***											
I	0.463	***0.789***										
P	0.122	0.302	***0.833***									
PL	0.174	0.271	0.285	***0.796***								
PS	0.401	0.335	0.372	0.252	***0.798***							
RO	0.427	0.350	0.390	0.258	0.553	***0.821***						
SL	0.115	0.267	0.528	0.565	0.392	0.375	***0.836***					
SW	0.311	0.454	0.443	0.359	0.485	0.455	0.388	***0.804***				
T	0.117	0.282	0.297	0.389	0.264	0.283	0.419	0.370	***0.774***			
TD	0.530	0.471	0.280	0.291	0.496	0.527	0.363	0.388	0.288	***0.807***		
V	0.492	0.321	0.063	0.249	0.252	0.297	0.148	0.229	0.082	0.443	***0.816***	
VC	0.455	0.331	0.213	0.281	0.416	0.513	0.269	0.369	0.152	0.530	0.569	***0.786***

Note. Bold and italics = Square-root of AVE.

Table 11. Discriminant Validity—Cross Loadings.

	BK	I	P	PL	PS	RO	SL	SW	T	TD	V	VC
BK1	0.739	0.370	0.057	0.169	0.285	0.283	0.061	0.187	0.144	0.377	0.317	0.293
BK2	0.882	0.438	0.079	0.142	0.326	0.349	0.056	0.216	0.065	0.419	0.399	0.364
BK3	0.838	0.397	0.080	0.145	0.351	0.319	0.086	0.270	0.099	0.413	0.442	0.411
BK4	0.786	0.310	0.105	0.183	0.358	0.332	0.153	0.292	0.099	0.404	0.393	0.387
BK5	0.877	0.399	0.108	0.139	0.345	0.407	0.108	0.251	0.117	0.486	0.471	0.397
BK6	0.790	0.364	0.140	0.098	0.314	0.389	0.093	0.296	0.060	0.489	0.390	0.377
I1	0.383	0.732	0.160	0.217	0.248	0.225	0.189	0.300	0.264	0.351	0.293	0.276
I2	0.314	0.820	0.280	0.228	0.239	0.220	0.242	0.375	0.246	0.345	0.232	0.250
I3	0.316	0.822	0.272	0.206	0.259	0.257	0.223	0.372	0.184	0.369	0.245	0.269
I4	0.386	0.783	0.238	0.248	0.302	0.296	0.237	0.349	0.182	0.357	0.242	0.276
I5	0.395	0.824	0.225	0.207	0.269	0.318	0.202	0.347	0.237	0.388	0.259	0.244
I6	0.390	0.751	0.247	0.185	0.269	0.326	0.176	0.398	0.228	0.410	0.247	0.251
P1	0.133	0.235	0.763	0.256	0.314	0.338	0.430	0.339	0.283	0.242	0.069	0.210
P2	0.098	0.260	0.887	0.275	0.338	0.313	0.474	0.394	0.262	0.242	0.017	0.162
P3	0.081	0.253	0.873	0.308	0.360	0.363	0.503	0.438	0.275	0.247	0.075	0.187
P4	0.078	0.207	0.826	0.211	0.259	0.328	0.418	0.310	0.165	0.184	0.060	0.180
P5	0.144	0.263	0.799	0.168	0.265	0.293	0.375	0.292	0.208	0.252	0.077	0.187
P6	0.057	0.275	0.845	0.193	0.309	0.307	0.432	0.437	0.279	0.221	0.014	0.132
PL1	0.186	0.224	0.228	0.767	0.234	0.212	0.467	0.310	0.384	0.249	0.244	0.248
PL2	0.165	0.210	0.249	0.861	0.229	0.197	0.440	0.342	0.319	0.219	0.184	0.226
PL3	0.088	0.146	0.198	0.852	0.181	0.165	0.419	0.265	0.301	0.178	0.176	0.169
PL4	0.150	0.276	0.214	0.726	0.173	0.197	0.441	0.218	0.205	0.245	0.261	0.249
PL6	0.089	0.197	0.237	0.764	0.182	0.241	0.454	0.282	0.325	0.242	0.107	0.201
PS1	0.251	0.212	0.264	0.235	0.760	0.391	0.326	0.351	0.213	0.374	0.170	0.319
PS2	0.322	0.265	0.333	0.224	0.870	0.466	0.363	0.409	0.216	0.428	0.208	0.365
PS3	0.352	0.309	0.334	0.249	0.863	0.464	0.333	0.403	0.244	0.432	0.258	0.392
PS4	0.330	0.243	0.262	0.244	0.757	0.423	0.314	0.367	0.182	0.355	0.171	0.318
PS5	0.285	0.241	0.279	0.104	0.756	0.434	0.273	0.382	0.197	0.374	0.141	0.276
PS6	0.380	0.330	0.304	0.152	0.774	0.459	0.263	0.408	0.208	0.403	0.247	0.313
RO1	0.333	0.266	0.270	0.207	0.404	0.779	0.317	0.346	0.202	0.369	0.224	0.422
RO2	0.333	0.288	0.324	0.231	0.419	0.875	0.310	0.382	0.243	0.435	0.246	0.457
RO3	0.315	0.300	0.344	0.231	0.476	0.856	0.346	0.407	0.211	0.436	0.237	0.441
RO4	0.391	0.296	0.314	0.246	0.465	0.799	0.321	0.368	0.249	0.448	0.288	0.443
RO5	0.356	0.280	0.326	0.169	0.457	0.795	0.272	0.357	0.250	0.446	0.232	0.389
RO6	0.362	0.283	0.336	0.179	0.493	0.815	0.275	0.371	0.232	0.449	0.234	0.375
SL1	0.119	0.256	0.406	0.511	0.334	0.334	0.819	0.322	0.360	0.319	0.175	0.304
SL2	0.071	0.231	0.469	0.513	0.330	0.281	0.886	0.352	0.381	0.268	0.097	0.235
SL3	0.077	0.225	0.468	0.476	0.340	0.292	0.896	0.328	0.341	0.281	0.106	0.237
SL4	0.145	0.240	0.444	0.502	0.334	0.336	0.839	0.297	0.324	0.315	0.143	0.217
SL5	0.072	0.161	0.407	0.376	0.311	0.306	0.742	0.292	0.327	0.303	0.126	0.161
SL6	0.083	0.227	0.463	0.448	0.315	0.328	0.830	0.347	0.366	0.323	0.082	0.188

(continued)

Table 11. (continued)

	BK	I	P	PL	PS	RO	SL	SW	T	TD	V	VC
SW1	0.283	0.346	0.353	0.281	0.415	0.395	0.277	0.772	0.363	0.309	0.159	0.327
SW2	0.216	0.364	0.374	0.335	0.357	0.324	0.356	0.837	0.323	0.302	0.161	0.245
SW3	0.201	0.367	0.329	0.291	0.371	0.339	0.282	0.836	0.287	0.274	0.161	0.280
SW4	0.279	0.336	0.333	0.300	0.412	0.378	0.319	0.789	0.263	0.335	0.219	0.333
SW5	0.292	0.406	0.363	0.258	0.395	0.401	0.319	0.791	0.293	0.351	0.244	0.337
SW6	0.203	0.361	0.403	0.263	0.383	0.342	0.309	0.799	0.257	0.288	0.146	0.238
T1	0.151	0.249	0.220	0.273	0.260	0.266	0.311	0.275	0.744	0.255	0.081	0.180
T2	0.059	0.204	0.237	0.333	0.242	0.215	0.375	0.324	0.829	0.230	0.039	0.097
T3	0.056	0.176	0.201	0.265	0.168	0.209	0.295	0.260	0.811	0.207	0.045	0.092
T6	0.086	0.247	0.266	0.331	0.122	0.168	0.310	0.284	0.707	0.190	0.088	0.083
TD1	0.385	0.357	0.237	0.235	0.423	0.447	0.317	0.316	0.231	0.804	0.334	0.458
TD2	0.446	0.386	0.260	0.271	0.432	0.477	0.318	0.318	0.233	0.848	0.387	0.456
TD3	0.441	0.434	0.240	0.246	0.416	0.455	0.327	0.320	0.241	0.867	0.405	0.457
TD4	0.413	0.343	0.228	0.272	0.444	0.354	0.309	0.349	0.267	0.748	0.300	0.397
TD5	0.438	0.400	0.198	0.239	0.369	0.408	0.288	0.311	0.248	0.830	0.390	0.404
TD6	0.436	0.351	0.176	0.135	0.311	0.397	0.175	0.261	0.175	0.738	0.322	0.387
V1	0.350	0.299	0.081	0.261	0.243	0.237	0.227	0.228	0.117	0.358	0.733	0.453
V2	0.449	0.261	0.030	0.171	0.199	0.228	0.077	0.163	0.000	0.360	0.873	0.451
V3	0.424	0.259	0.032	0.208	0.190	0.228	0.096	0.166	0.061	0.387	0.894	0.476
V6	0.379	0.229	0.060	0.183	0.192	0.281	0.077	0.186	0.092	0.338	0.753	0.472
VC1	0.316	0.266	0.122	0.231	0.314	0.399	0.213	0.270	0.149	0.414	0.425	0.785
VC2	0.394	0.268	0.140	0.219	0.318	0.425	0.177	0.269	0.133	0.412	0.442	0.830
VC3	0.373	0.229	0.153	0.245	0.338	0.425	0.216	0.250	0.060	0.434	0.480	0.866
VC4	0.303	0.228	0.182	0.198	0.356	0.346	0.207	0.289	0.045	0.359	0.391	0.705
VC6	0.389	0.305	0.228	0.212	0.316	0.417	0.240	0.361	0.196	0.453	0.486	0.735

Note. Grey highlighting illustrates that Cross-loadings indicated that each indicator loaded strongly onto its parent construct and not on other constructs.

Table 12. Discriminant Validity—HTMT.

	BK	I	P	PL	PS	RO	SL	SW	T	TD	V
BK											
I	0.519										
P	0.132	0.335									
PL	0.196	0.306	0.317								
PS	0.450	0.378	0.410	0.287							
RO	0.469	0.389	0.429	0.289	0.617						
SL	0.125	0.299	0.578	0.632	0.436	0.412					
SW	0.341	0.511	0.490	0.409	0.545	0.504	0.428				
T	0.142	0.341	0.351	0.475	0.308	0.330	0.495	0.443			
TD	0.589	0.529	0.307	0.325	0.556	0.584	0.397	0.432	0.342		
V	0.568	0.377	0.080	0.292	0.293	0.345	0.169	0.264	0.111	0.514	
VC	0.518	0.384	0.242	0.324	0.482	0.587	0.305	0.421	0.184	0.609	0.679

Table 13. Variance Inflation Factor.

Constructs	VIF	Constructs	VIF	Constructs	VIF
BK	1.758	RO	1.882	VC	1.922
I	1.549	SL	2.044	PL	1.641
P	1.629	SW	1.758	T	1.364
PS	1.748	V	1.707		

between 11 independent variables (T, V, VC, SL, RO, PS, P, SW, I, BK) and the dependent variable (TD). This was done to address the first research question: What factors do undergraduate EFL learners perceive as affecting the text difficulty of model paragraphs.

The path relationships were examined using the PLS-SEM two-tailed bootstrap procedure (5,000 samples)

Table 14. Direct Relationship Results.

	Path	β	*T*	*p*	Supported
H1	T > TD	.08	2.02	.04	Yes
H2	PL > TD	−.02	0.47	.64	No
H3	V > TD	.11	2.25	.02	Yes
H4	VC > TD	.17	3.04	.02	Yes
H5	SL > TD	.12	2.29	.02	Yes
H6	RO > TD	.15	2.61	.01	Yes
H7	PS > TD	.14	2.63	.01	Yes
H8	SW > TD	−.03	0.61	.54	No
H9	P > TD	−.01	0.29	.77	No
H10	I > TD	.16	3.17	.00	Yes
H11	BK > TD	.20	3.95	.00	Yes

Table 15. Sub Hypotheses $H_{1a\text{-}d.}$

H	Path	β	*T*	*p*	Supported	Type
$H_{1a:}$	RO > T > TD	.01	1.56	.17	No	No mediation
H_{1b}	PS > T > TD	.01	1.41	.16	No	No mediation
H_{1c}	I > T > TD	.02	1.74	.08	No	No mediation
H_{1d}	BK > T > TD	−.01	1.50	.13	No	No mediation

using three measures: path coefficients (β), *T* values, and *p* values. The results showed that eight hypotheses were supported (H_1, H_3, H_4, H_5, H_6, H_7, H_{10}, H_{11}), indicating significant direct relationships between eight independent variables (T, V, VC, SL, RO, PS, I, BK) and the dependent variable TD at below 0.05 *p* level and above the 1.96 T threshold.

An examination of the comparative strength of these relationships showed BK had the highest path coefficient, followed by VC, I, RO, PS, SL, V, and T. No significant effect was found for three variables (PL, P, SW). Thus, three hypotheses (H_2, H_8, H_9) were not supported (Table 14).

Mediation Analysis

Mediation was explored to address the second research question: What mediating factor relationships do undergraduate EFL learners perceive as affecting the text difficulty of model paragraphs? Mediation was explored by first assessing the indirect effect of the independent variable through the mediating variable to the independent variable (X > M > Y) (Nitzl et al., 2022). If significant, mediation was investigated further. Where the direct effect between the independent variable and the dependent variable was significant (X > Y), partial mediation was identified, as both the mediating (M) and independent path are contributing. In cases where the direct path was insignificant, full mediation was identified, as only the mediating variable (M) had an effect.

Seven sets of sub hypotheses were explored, 26 sub hypotheses in total. These explored how seven variables (T, PL, V, RO, SW, P, BK) acted as mediating variables between seven other variables (VC, RO, PS, P, SL, I, BK) and the dependent variable TD. The significance of the mediating relationships was assessed using a two-tailed bootstrapping (5,000 resamples) and the resulting three measures: path coefficients (β), *T* values, and *p* values. The results showed that seven of the 26 sub hypotheses were supported.

$H_{1a\text{-}d}$ explored the mediating relationships between T and four variables (PS, RO, I, BK) and TD. That is, whether T mediates the relationship between these variables and participants' perceptions of TD. $H_{1a\text{-}d}$ were not supported. It was found that T did not mediate the relationship between each of the variables (RO, PS, I, BK) and TD. That is, the mediating paths (RO > T > TD; PS > T > TD; I > T > TD; BK > T > TD, respectively) were not significant (Table 15). Hence, no mediating relationships were identified.

$H_{2a\text{-}e}$ explored the mediating relationships between five variables (V, VC, PS, I, BK) and PL. That is, how PL mediates the relationship between these variables and participants' perceptions of TD. $H2_{a\text{-}e}$ were not supported as the results showed that PL did not significantly mediate the relationship between the five variables (V, VC, PL, I, K) and TD. That is, the mediation paths (V > PL > TD; VC > PL > TD; PS > PL > TD; >PL > I > TD; BK > PL > TD, respectively) were not significant (Table 16).

Table 16. Sub Hypotheses $H_{2a\text{-}e.}$

H	Path	β	*T*	*p*	Supported	Type
H2a	V > PL > TD	−.003	0.41	.68	No	No mediation
H2b	VC > PL > TD	−.003	0.41	.68	No	No mediation
H2c	PS > PL > TD	−.003	0.42	.68	No	No mediation
H2d	I > PL > TD	.004	0.45	.66	No	No mediation
H2e	BK > PL > TD	.022	0.37	.71	No	No mediation

Table 17. Sub Hypotheses $H_{3a\text{-}d.}$

H	Path	β	*T*	*p*	Supported	Type
H_{3a}	VC > V > TD	.059	2.14	.03	Yes	Partial mediation
H_{3b}	SL > V > TD	−.001	0.62	.53	No	No mediation
H_{3c}	RO > V > TD	−.001	0.34	.73	No	No mediation
H_{3d}	I > V > TD	.019	1.91	.06	No	No mediation

Table 18. Sub Hypotheses $H_{5a\text{-}d.}$

H	Path	β	*T*	*p*	Supported	Type
H_{5a}	RO > SL > TD	.01	1.40	.16	No	No mediation
H_{5b}	PS > SL > TD	.25	1.76	.08	No	No mediation
H_{5c}	P > SL > TD	.05	2.21	.03	Yes	Full mediation
H_{5d}	I > SL > TD	.01	0.85	.40	No	No mediation

Table 19. Sup Hypothesis $H_{6a.}$

H	Path	β	*T*	*p*	Supported	Type
H_{6a}	VC > RO > TD	.08	2.49	.01	Yes	Partial mediation

$H_{3a\text{-}d}$ explored the mediating relationships for four variables (VC, SL, RO, I) and V. That is, how V mediates the relationship between these variables and participants' perceptions of TD. H_{3a} was supported. A partial mediating relationship was identified for variable VC, as the mediating path (VC > V > TD) and the direct path between VC and TD (VC > TD) were found to be significant. No mediation relationships were found between the other three variables (SL, RO, I) and variables V and TD, as the mediation paths (SL > V > TD; RO > V > TD; I > V > TD, respectively) were not significant (Table 17).

$H_{5a\text{-}d}$ explored the mediating relationship between four variables (RO, PS, P, I), SL, and TD, that is, how SL mediates the relationship between these variables and participants' perceptions of TD. $H5_a$, $H5_b$, and $H5_d$ were not supported. For these three variables (RO, PS, I), no mediation relationships were found as their mediating paths (RO > SL > TD; PL > PL > TD; I > SL > TD) were not significant. However, $H5_c$ was supported, as mediating path (P > SL > TD) was significant, and the direct path (P > TD) was not significant. Hence, full mediation was identified (Table 18).

Sub-hypothesis H_{6a} explored the mediating relationship for one variable (VC) and how RO mediates the relationship between VC and participants' perceptions of TD. H_{6a} was supported. A partial mediation relationship was identified for RO as both its mediating relationship (VC > RO > TD) and the direct relationship between VC and TD (VC > TD) were significant (Table 19).

$H_{8a\text{-}b}$ explored the mediating relationship for two variables (RO, PS) and how SWs mediate the relationships between these variables and participants' perceptions of text difficulty. $H_{8a\text{-}b}$ were not supported. No mediation

Table 20. Sub Hypotheses $H_{8a\text{-}b.}$

H	Path	β	*T*	*p*	Supported	Type
H_{8a}	RO > SW > TD	−.01	0.60	.55	No	No mediation
H_{8b}	PS > SW > TD	−.01	0.60	.56	No	No mediation

Table 21. Sub Hypotheses $H_{11a\text{-}f.}$

H	Path	β	*T*	*P*	Supported	Type
H_{11a}	V > BK > TD	.05	2.97	.00	Yes	Partial mediation
H_{11b}	VC > BK > TD	.02	1.25	.22	No	No mediation
H_{11c}	RO > BK > TD	.03	1.92	.06	No	No mediation
H_{11d}	PS > BK > TD	.03	2.42	.02	Yes	Partial mediation
H_{11e}	P > BK > TD	−.02	2.24	.03	Yes	Full mediation
H_{11f}	I > BK > TD	.05	3.12	.02	Yes	Partial mediation

relationships were identified for either variable, as the path relationships (RO > SW > TD; PS > SW > TD) were not significant (Table 20).

$H_{11a\text{-}f}$ explored the mediating relationships of six variables (V, VC, RO, PS, P, I) and how BK mediates the relationships between these variables and participants' perceptions of TD. H_{11a}, H_{11d}, and H_{11f} were supported. Three variables (V, PS, I) showed partial mediation relationships, since the mediating paths (V > BK > TD; PS > BK > TD; I > BK > TD) and the direct paths (V > TD; PS > TD; I > TD) were significant (Table 21). H_{11e} was also supported. A full mediation relationship was identified for one variable (P) as the mediating path (P > BK > TD) was significant, but the direct path (P > TD) was not significant. $H_{11b\text{-}c}$ were not supported. No mediation relationships were identified for two variables (VC, RO) as the mediation paths were not significant (VC > BK > TD; RO > BK > TD).

Explanatory Power: Coefficients of Determination (R^2)

The explanatory power of the PLS-SEM model was explained using R^2 in the context of the study (Sarstedt et al., 2017), as the range of weak to substantial has been described differently in different fields: 0.10, for stock returns (Raithel et al., 2012); marketing, 0.25, 0.50, 0.75 (weak, moderate, substantial) (Hair et al., 2011); behavioral sciences, a range of 0.02 to 0.26 ($R^2 < 0.02$, very weak; $0.02 \leqslant R^2 < 0.13$, weak; $0.13 \leqslant R^2 < 0.26$, moderate; $R^2 \geqslant 0.26$, substantial) (Cohen, 1988). For attitudinal research, as in this study, the following has been suggested: $R^2 < 0.19$, very weak; $0.19 \leqslant R^2 < 0.33$, weak; $0.33 \leqslant R^2 < 0.67$, moderate; $R^2 \geqslant 0.67$, substantial (Chin 1998).

The explanatory power of the proposed PLS-SEM predictive model was found to be 0.508 R^2, a moderate to substantial measure (Chin 1998) (Figure 1).

Predictive Relevance (Q^2)

To determine predictive relevance (Q^2), PLS-SEM's blindfolding technique (Chin, 2010) was employed according to the rule of thumb that values larger than zero represent relevance. It was found that the model had a 0.351 measure, indicating a moderate to high level of predictive relevance.

Discussion

A review of the extant literature showed that previous investigations had only explored one or a small number of variables and a limited number of mediating relationships. Moreover, what primary and mediating features EFL learners perceive as contributing to the difficulty of model paragraphs was found to be noticeably absent. To address this, this study investigated what factors undergraduate EFL learners perceive affecting the difficulty of model paragraphs excerpted from writing coursebooks. Two research questions were set to explore this, with hypotheses related to each. To assess the results, a PLS-SEM model was posed and tested.

Regarding RQ1 (What factors do undergraduate EFL learners perceive as affecting the text difficulty of model paragraphs), the findings showed that 8 of the 11 direct hypotheses (H_1, H_3, H_4, H_5, H_6, H_7, H_{10}, H_{11}) were supported; that is, eight independent variables (titles, vocabulary, vocabulary in context, sentence length, rhetorical organization, paragraph structure, interest, background

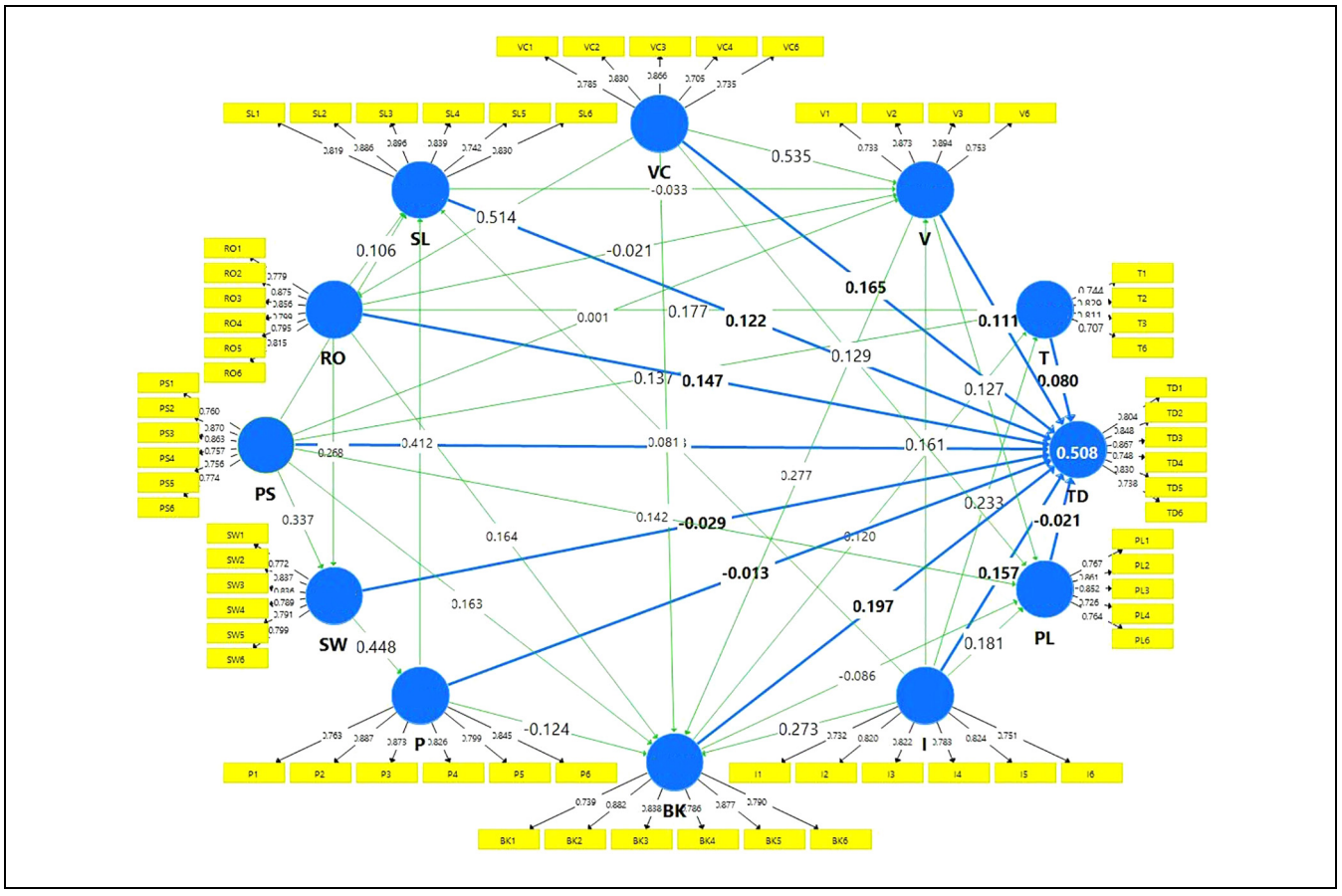

Figure 1. Explanatory power: coefficients of determination (R^2).
Note. $R^2 = 0.508$.

knowledge) significantly affected students' perceptions of text difficulty. Conversely, three hypotheses were not supported (H_2, H_8, H_9). That is, three variables (paragraph length, signal words, punctuation) were not found to have a significant effect.

Regarding RQ2 (What mediating factor relationships do undergraduate EFL learners perceive as affecting the text difficulty of model paragraphs?), seven mediating hypotheses were supported, while 19 were not. That is, four variables were identified as significant mediating variables (vocabulary, sentence length, rhetorical organization, background knowledge), and three were not (titles, paragraph length, signal words).

The PLS-SEM model accounted for 0.508 R^2 of undergraduate EFL learners' perceptions of what factors contribute to the readability of model paragraphs and was found to have high predictive relevance, 0.315 Q^2.

The findings corroborate, contradict, and further the extant literature. These are organized according to the research questions foci (i.e., direct, mediating relationships), respectively. Titles, for instance, demonstrated a significant relationship with text difficulty, consistent with previous studies (Bartlett, 1932; Carrell, 1983; Dooling & Lachman, 1971; Noor, 2006). Nevertheless, no significant mediating relationships were shown with other variables (rhetorical organization, paragraph structure, interest, background knowledge), contrary to literature that has demonstrated such effects (Ahmadi, 2011; Baker, 2020; Bartlett, 1932; Bock, 1980).

Paragraph length showed no significant direct relationship with text difficulty, a surprising finding contrary to similar research (Freedle & Kostin, 1991, 1992, 1993; Gopal & Mahmud, 2019; Moon, 2019) but consistent with research showing no relationship (Jalilehvand, 2012; Lee, 1999; Mehrpour & Riazi, 2004). Similarly, no mediating relationships were found, contrary to studies that have illustrated paragraph length's relationship with other variables (vocabulary, vocabulary in context, paragraph structure, background knowledge) (Baker, 2020; Bock, 1980; Keenan et al., 1985; Mandler & Johnson, 1977; Reder & Anderson, 1980; Shokouhi & Askari, 2010). These findings may be attributable to the similarity in the lengths of the selected paragraphs.

Vocabulary was found to have a significant direct relationship with text difficulty, corroborating literature with similar findings (Baker, 2020; Chou, 2011; Kameli &

Baki, 2013; Qian, 2002; Salyer, 1990; Yorio, 1971). Furthermore, a mediating relationship with vocabulary in context was found, consistent with literature indicating such relationships (Baker, 2020, 2021; Guo, 2008; Haynes & Baker, 1993). However, mediating relationships with other variables were not found (sentence length, rhetorical organization, interest), contrary to other works (Alkhaleefah, 2017; Baker, 2020; Carrell, 1983; Guo, 2008).

Vocabulary in context was found to have a significant direct relationship with text difficulty, supporting work that has shown that vocabulary in context clues can affect text difficulty (Ahmad et al., 2018; Bengeleil & Paribakht, 2004; Cooper, 1999; Dubin & Olshtain, 1993; Haynes, 1993).

Sentence length showed a relationship with text difficulty, which is consistent with previous work that has demonstrated this relationship (Coleman, 1962; Coleman & Miller, 1968; Freedle & Kostin, 1992; Glazer, 1974; McElree, 2000; McLaughlin, 1969; Mikk, 2008; Nilagupta, 1977). Contrary to previous research (McElree, 2000; McLaughlin, 1969; Mikk & Kukemelk, 2010), no mediating relationships with rhetorical organization, paragraph structure, or interest were found. However, a relationship with punctuation was identified, supporting literature showing that sentences become more difficult as they become more complex (Coleman, 1962; Coleman & Miller, 1968; Glazer, 1974; Sherman, 1893). That is, longer sentences tend to have a more extensive variety of punctuation (Baker, 2020), and students' understanding can play a part (Durkee, 1952).

It was found that rhetorical organization and text complexity were associated, which is similar to taxonomy research showing that some texts are more complex than others (Alkhaleefah, 2017; Amiri et al., 2012; Baker, 2021; Carrell, 1984a; Freedle & Kostin, 1991, 1993; Meyer & Freedle, 1984; Putra, 2012; Saadatnia et al., 2016). A mediating relationship with vocabulary in context was also identified, supporting work showing that rhetorical organization assists students in guessing clues provided by vocabulary in context (Baker, 2020).

Paragraph structure was found to be associated with text difficulty, supporting research showing students consider how well the text is structured when attempting to understand it (Baker, 2020; Carrell, 1984a, 1984b; Kieras, 1978; Lorch & Lorch, 1985; Ritzer, 1994; Thorndyke, 1975). This is also consistent with work that has illustrated that students' awareness of structure plays an important role (Namjoo & Marzban, 2012; Shemshadsara et al., 2019).

Signal words did not significantly affect text difficulty, which is in alignment with research that has found that signal words do not affect students' reading comprehension (Meyer, 1975), but not with research that argues the importance of signal words (Baker, 2021; Van Silfhout et al., 2014) or research that states that signal words affect students' understanding of texts (Aidinlou & Pandian, 2011; Al-Surmi, 2011; Miccinati, 1975; Roen, 1984; Xu et al., 2019). Furthermore, signal words did not mediate the relationship between rhetorical organization and paragraph structure, contrary to previous literature that argues signal words assist readers in forming a mental representation of the relationships within a text (Xu et al., 2019). This surprising result may be attributable to the limited number of signal words presented in the texts.

Punctuation was not found to have a direct relationship with text difficulty, contrary to research that has illustrated the importance of punctuation (Alsubaie, 2014; Backscheider, 1972; Carr, 1978; Carver, 1970; Neff, 1932; Pathan & Al-Dersi, 2013; Suliman et al., 2019). It was found, however, to be related to sentence length (see above discussion).

Interest was found to have a significant direct relationship with students' perceptions of text difficulty, supporting work that has demonstrated interest can impact understanding (Erçetin, 2010). Interest was also shown to have a relationship with other features, but one that was mediated by another feature (see background knowledge).

Background knowledge was found to significantly influence text difficulty, confirming previous research suggesting background knowledge influences readers' comprehension (Carrell, 1983; Chau et al., 2019; Florencio, 2004; Ghorbandordinejad & Bayat, 2014; Khataee & Davoudi, 2018; Nelson, 1987; Steffensen et al., 1979). Background knowledge was also found to have a mediating relationship with several other features (vocabulary, paragraph structure, punctuation, and interest) (Ay & Bartan, 2012; Bugel & Buunk, 1996; Carrell, 1987; Carrell & Wise, 1998; Demir, 2012; Kelsen, 2016; Sheridan et al., 2019). However, no relationship with vocabulary in context or rhetorical organization was found, which is in contrast to Carrell (1983), who argued that background knowledge facilitates inference and can compensate for difficulties with rhetorical organization.

Conclusion

Overall, the study identified eight features that had a significant direct effect on students' perceptions of difficulty and four mediating variables. Considering these features and relationships as individual hypotheses in isolation, as was done in previous studies, is insightful. However, this investigation, through the unique methodological lens of PLS-SEM, offers a unique contribution to the literature as the resulting complex statistical model provides a

holistic explanation of the direct and mediating relationships (Hair et al., 2018; Wong, 2019), which accounted for a moderate to substantial measure (Chin 1998; Hair et al., 2021) of students' perceptions regarding what features contribute to the difficulty of model paragraphs with a moderate to high predictive relevance.

We hope this model will be useful to various stakeholders, as paragraph difficulty (text readability) is an important but much-neglected topic. These findings are potentially useful to teachers and syllabus designers who wish to understand the appropriateness of materials for learners but might otherwise make decisions on intuitive grounds (Fulcher, 1997), resulting in choosing materials that are too difficult and thus damaging to learning and motivation. Members of the publishing industry and material writers may also find the results useful in creating appropriate-level model paragraphs. Additionally, the results have the potential to inform the research community and further the literature by extending our knowledge of what contributes to readability.

Implications and Suggestions for Future Study

The results of this study (the list of features and their direct and mediating relationships presented and accompanying PLS-SEM model) have practical implications, as they have the potential to be of use to teachers of writing, material writers, and members of the publishing industry and further readability literature. However, they beg several questions that might be explored in future studies.

First, the model explained 50.8% ($R^2 = 0.508$) of undergraduate EFL learners' perceptions of text difficulty of model paragraphs, a moderate to high result. Nevertheless, the question of what other variables or mediating relationships might account for a larger percentage needs to be addressed. As this study adopted a transmittal approach, testing path relationships based on existing literature (Nitzl et al., 2022), an alternative segmentation approach of exploring additional paths and testing a contrasting model could add to existing theory and explain a larger portion of R^2.

Second, the Q^2 result of 0.351 indicates that this model has high predictive relevance and may be generalizable beyond the study's context; nevertheless, replicability and generalizability are important concerns (National Academies of Sciences, Engineering, & Medicine, 2019; Strube, 2000). Accepting this, this exploration was regionally (i.e., Vietnam), contextually (undergraduate English majors), and material (one text) specific. Thus, additional studies with various contexts and materials would be prudent to gain insight into additional experiences. Third, although PLS-SEM was tested in 1982 (Wold, 1982), serves as a standard research methodology (American Psychological Association [APA], 2021a, 2021b), is widely used (Hair et al., 2021), and is increasingly used in education, there is a notable absence of PLS-SEM models that have explored this area. It is therefore hoped that this study will serve as a foundation for future investigations and discussions of the features that affect the readability of model texts.

Acknowledgments

We would like to thank the Head Editor, editors, and reviewers of the Sage Open Journal for their suggestions and guidance, Dr. James Gaskin for his expertise with PLS-SEM, Katherine Kurowski for her APA reference suggestions, and Luu Thi Thanh An and Thanh Nguyen for their help as research assistants.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Request for funding has been sent to Ton Duc Thang University, Ho Chi Minh City, Vietnam.

ORCID iDs

John R. Baker https://orcid.org/0000-0003-3379-4751
Thao Quang Le https://orcid.org/0000-0002-5083-5730

Supplemental Material

Supplemental material for this article is available online.

Data Availability Statement

Questions regarding this research can be sent to the corresponding author, Dr. John R. Baker, Creative Language Center, Ton Duc Thang University, 19 Nguyen Huu Tho St, Tan Phong Ward, Dist. 7, Ho Chi Minh City, Vietnam. Email: drjohnrbaker@tdtu.edu.vn.

References

Abbott, M. L. (2006). ESL reading strategies: Differences in Arabic and Mandarin speaker test performance. *Language Learning*, *56*(4), 633–670. https://doi.org/10.1111/j.1467-9922.2006.00391.x

Ahmadi, A. (2011). Comprehension of a non-text: The effect of the title and ambiguity tolerance. *Journal of Pan-Pacific Association of Applied Linguistics*, *15*(1), 163–176.

Ahmad, S. N., Muhammad, A. M., & Kasim, A. A. M. (2018). Contextual clues vocabulary strategies choice among business management students. *English Language*

Teaching, *11*(4), 107–116. https://doi.org/10.5539/elt.v11n4p107

Aidinlou, N. A., & Pandian, A. A. (2011). The impact of local and global conjunctions on ESL reading comprehension: A systemic perspective. *Journal of Language Teaching and Research*, *2*(2), 387–395. https://doi.org/10.4304/jltr.2.2.387-395

Alkhaleefah, T. A. (2017). Saudi EFL learners' reported reading problems and strategic processing of text types: A think-aloud case study. *Reading Psychology*, *38*(7), 687–730. https://doi.org/10.1080/02702711.2017.1336660

Allington, R. (2002). You can't learn much from books you can't read. *Educational Leadership*, *60*(3), 16–19.

Alsubaie, M. A. A. (2014). *An exploration of reading comprehension challenges in Saudi Arabian University EFL students* [Doctoral dissertation, The University of Exeter]. Open Research Exeter. http://hdl.handle.net/10871/15981

Al-Surmi, M. (2011). Discourse markers and reading comprehension: Is there an effect? *Theory and Practice in Language Studies*, *1*(12), 1673–1678. https://doi.org/10.4304/tpls.1.12.1673-1678

American Psychological Association. (2021b). *Journal article reporting standards*. https://apastyle.apa.org/jars

American Psychological Association. (Eds.) (2021a). *Publication manual of the American psychological association*. Author.

Ames, W. S. (1966). The development of a classification scheme of contextual aids. *Reading Research Quarterly*, *2*(1), 57–82. https://doi.org/10.2307/747039

Amiri, F., Zainal, Z., & Samad, A. A. (2012). Effects of text structure on the Iranian students' reading comprehension performance. *Procedia - Social and Behavioral Sciences*, *66*, 402–409. https://doi.org/10.1016/j.sbspro.2012.11.283

Atamturk, N. (2018). Graduate students' views of topic attributes in e-books. *IIOABJ*, *9*(3), 130–134.

Ay, S., & Bartan, Ö. Ş. (2012). The effect of topic interest and gender on reading test types in a second language. *The Reading Matrix*, *12*(1), 62–79.

Backscheider, P. (1972). Punctuation for the reader–A teaching approach. *The English Journal*, *61*(6), 874. https://doi.org/10.2307/813995

Bae, M., & Lee, B. (2018). Effects of text length and question type on test-takers' performance on fill-in-the-blank items in Korean CSAT. *English Teaching*, *73*(4), 149–174. https://doi.org/10.15858/engtea.73.4.201812.149

Baker, J. R. (2019). Writing about the writing center in the Asian context: Exploring the mis/match between the reading levels of self-access materials and the students who visit the center. *The Asian ESP Journal*, *15*(3), 256–285.

Baker, J. R. (2020). A checklist for use with the Lexile readability formula when choosing materials for writing center self-access libraries. *The Asian ESP Journal*, *16*(4), 9–68.

Baker, J. R. (2021). Exploring how rhetorical organization contributes to the readability of essays. *Journal of Language and Education*, *7*(2), 78–92. https://doi.org/10.17323/jle.2021.11240

Baker, J. R. (2022). Investigating the effects of signal words on the readability of writing centre self-access library materials. *The Language Learning Journal*, *51*(2), 66–76. https://doi.org/10.1080/09571736.2022.2099958

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Bengeleil, N., & Paribakht, T. (2004). L2 reading proficiency and lexical inferencing by university EFL learners. *The Canadian Modern Language Review/La revue canadienne des langues vivantes*, *61*(2), 225–250. https://doi.org/10.3138/cmlr.61.2.225

Benitez-Rivera, W. I. (2013). *Efficacy of attention to commas (A2C) strategy for sentence comprehension in English language learners (ELLs) (Publication No. 3592946)* [Doctoral dissertation. Howard University]. ProQuest Dissertations & Theses Global.

Bock, M. (1980). Some effects of titles on building and recalling text structures. *Discourse Processes*, *3*(4), 301–311. https://doi.org/10.1080/01638538009544494

Bransford, J. D., & Johnson, M. K. (1973). Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing* (pp. 383–439). Academic Press.

Britton, B. K., & Black, J. B. (2017). Understanding expository text: From structure to process and world knowledge. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text: A theoretical and practical handbook for analyzing explanatory text* (pp. 1–9). Routledge.

Brown, J. D. (2001). *Using surveys in language programs*. Cambridge University Press.

Bugel, K., & Buunk, B. P. (1996). Sex differences in foreign language text comprehension: The role of interests and prior knowledge. *Modern Language Journal*, *80*(1), 15–31. https://doi.org/10.2307/329055

Carr, M. N. (1978). *An instrument to assess readers' ability to interpret internal punctuation (Publication No. 7823171)* [Doctoral Dissertation, Arizona State University]. ProQuest Dissertations & Theses Global.

Carrell, P. L. (1983). Background knowledge in second language comprehension. *Language learning and Communication*, *2*(1), 25–34.

Carrell, P. L. (1984a). The effects of rhetorical organization on ESL readers. *TESOL Quarterly*, *18*(3), 441–469. https://doi.org/10.2307/3586714

Carrell, P. L. (1984b). Evidence of a formal schema in second language comprehension. *Language Learning*, *34*(2), 87–108. https://doi.org/10.1111/j.1467-1770.1984.tb01005.x

Carrell, P. L. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly*, *21*(3), 461–481. https://doi.org/10.2307/3586498

Carrell, P. L. (1992). Awareness of text structure: Effects on recall. *Language Learning*, *42*(1), 1–18. https://doi.org/10.1111/j.1467-1770.1992.tb00698.x

Carrell, P. L., & Wise, T. E. (1998). The relationship between prior knowledge and topic interest in second language reading. *Studies in Second Language Acquisition*, *20*(3), 285–309. https://doi.org/10.1017/s0272263198003015

Carver, R. P. (1970). Effect of a "chunked" typography on reading rate and comprehension. *E-Journal of Applied Psychology*, *54*(3), 288–296. https://doi.org/10.1037/h0029266

Chall, J. S., Bissex, G. L., Conrad, S. S., & Harris-Sharples, S. (1996). *Qualitative assessment of text difficulty*. Brookline Books.

Chall, J. S., & Dale, E. (1995). *Readability revisited: The new Dale-Chall readability formula*. Brookline Books.

Chau, N. N. B., Liu, S., & Lopez, K. (2019). Effect of cultural familiarity on reading comprehension performance: A case-study of Vietnamese and Chilean EFL learners. *Tạp Chí Khoa Học Ngôn Ngữ Và Văn Hóa* , *3*(1), 2–10.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Lawrence Erlbaum.

Chin, W. W. (2010). How to write up and report PLS analyses. In V. E. Vinzi, W. W. Chin, J. Jörg Henseler & H. Wang (Eds.), *Handbook of partial least squares* (pp. 655–690). Springer.

Chou, P. T. M. (2011). The effects of vocabulary knowledge and background knowledge on reading comprehension of Taiwanese EFL students. *Electronic Journal of Foreign Language Teaching*, *8*(1), 108–115.

Chung, J. S. L. (2000). Signals and reading comprehension — Theory and practice. *System*, *28*(2), 247–259. https://doi.org/10.1016/s0346-251x(00)00010-5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Erlbaum.

Coleman, E. B. (1962). Improving comprehensibility by shortening sentences. *E-Journal of Applied Psychology*, *46*(2), 131–134. https://doi.org/10.1037/h0039740

Coleman, E. B., & Miller, G. R. (1968). A measure of information gained during prose learning. *Reading Research Quarterly*, *3*(3), 369–386. https://doi.org/10.2307/747010

Commander, N. E., & Stanwyck, D. J. (1997). Illusion of knowing in adult readers: Effects of reading skill and passage length. *Contemporary Educational Psychology*, *22*(1), 39–52. https://doi.org/10.1006/ceps.1997.0925

Cooper, T. C. (1999). Processing of idioms by L2 learners of English. *TESOL Quarterly*, *33*(2), 233–262. https://doi.org/10.2307/3587719

Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, 34(2), 213–238. https://doi.org/10.2307/3587951

Davis, F. B. (1944). Fundamental factors of comprehension in reading. *Psychometrika*, *9*(3), 185–197. https://doi.org/10.1007/bf02288722

Davoudi, M., & Nafchi, A. M. (2016). The effect of lexical inferencing on the vocabulary learning and reading comprehension of Iranian intermediate EFL learners. *Modern Journal Of Language Teaching Methods*, *6*(3), 176.

Demir, Y. (2012). The effect of background knowledge and cultural nativization on reading comprehension and vocabulary inference. *Journal of Educational and Instructional Studies in the World*, *2*(4), 188–198.

Dewey, J. (1913). *Interest and effort in education*. Houghton Mifflin.

Dooling, D. J., & Lachman, R. (1971). Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, *88*(2), 216–222. https://doi.org/10.1037/h0030904

Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.

DuBay, W. H. (2007a). *Unlocking language: The classic readability studies*. Impact Information.

DuBay, W. H. (2007b). *Smart language: Readers, readability, and the grading of text*. Impact Information.

Dubin, F., & Olshtain, E. (1993). Predicting word meanings from contextual clues: Evidence from L1 readers. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 181–202). Ablex.

Durkee, F. M. (1952). Freshman reading problem: A proposed attack. *College English*, *14*(1), 30–33. http://www.jstor.org/stable/371826

Dwaik, R. A. (1997). *The role of lexical and syntactic knowledge in English as a foreign language reading comprehension (Publication No. 9731615)* [Doctoral dissertation, The Ohio State University]. ProQuest Dissertations & Theses Global.

Earle, J. (1890). *English prose its elements, history, and usage*. Smith, Elder & Co.

Erçetin, G. (2010). Effects of topic interest and prior knowledge on text recall and annotation use in reading a hypermedia text in the L2. *ReCALL*, *22*(2), 228–246. https://doi.org/10.1017/s0958344010000091

Flick, W. C., & Anderson, J. I. (1980). Rhetorical difficulty in scientific English: A study in reading comprehension. *TESOL Quarterly*, *14*(3), 345–351. https://doi.org/10.2307/3586599

Florencio, D. C. (2004). *The role of prior background knowledge in the reading comprehension of EFL Brazilian college students and American college students (Publication No. 3140017)* [Doctoral dissertation, The Pennsylvania State University]. ProQuest Dissertations & Theses Global.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.2307/3151312

Freedle, R., & Kostin, I. (1991). The prediction of SAT reading comprehension item difficulty for expository prose passages. *ETS Research Report Series*, *1991*(1), i-52. https://doi.org/10.1002/j.2333-8504.1991.tb01396.x

Freedle, R., & Kostin, I. (1992). The prediction of GRE reading comprehension item difficulty for expository prose passages for each of three item types: Main ideas, inferences and explicit statements. *ETS Research Report Series*, *1991*(2), i-53. https://doi.org/10.1002/j.2333-8504.1991.tb01426.x

Freedle, R., & Kostin, I. (1993). The prediction of TOEFL reading comprehension item difficulty for expository prose passages for three item types: Main idea, inference, and supporting idea items. *ETS Research Report Series*, *1993*(1), i–48. https://doi.org/10.1002/j.2333-8504.1993.tb01524.x

Fry, E. (2002). Readability versus Leveling. The Reading Teacher, *56*(3), 286–291. http://www.jstor.org/stable/20205195

Fulcher, G. (1997). Text difficulty and accessibility: Reading formulae and expert judgement. *System*, *25*(4), 497–513. https://doi.org/10.1016/S0346-251X(97)00048-1

Ghorbandordinejad, F., & Bayat, Z. (2014). The effect of cross-cultural background knowledge instruction on Iranian EFL learners' reading comprehension ability. *Theory and Practice in Language Studies*, *4*(11), 2373–2383. https://doi.org/10.4304/tpls.4.11.2373-2383

Gilliand, G. (1972). *Readability*. Hodder and Stoughton.

Glazer, S. M. (1974). Is sentence length a valid measure of difficulty in readability formulae? *The Reading Teacher*, *27*(5), 464–468. http://www.jstor.org/stable/20193535

Goh, S. T. (1990). The effects of rhetorical organization on expository prose on ESL readers in Singapore. *RELC Journal*, *21*(2), 1–11. https://doi.org/10.1177/003368290021002001

Gopal, R., & Mahmud, C. T. (2019). Prose reading: The influence of text-reader factors. *Studies in English Language and Education*, *6*(2), 187–198.

Gunning, T. G. (2003). The role of readability in today's classrooms. *Topics in Language Disorders*, *23*(3), 175–189.

Guo, Y. (2008). *The role of vocabulary knowledge, syntactic awareness and metacognitive awareness in reading comprehension of adult English language learners (Publication No. 3340717)* [Doctoral dissertation, Florida State University]. ProQuest Dissertations & Theses Global.

Ha, H. T. (2021). Exploring the relationships between various dimensions of receptive vocabulary knowledge and L2 listening and reading comprehension. *Language Testing in Asia*, *11*(1), 1–20. https://doi.org/10.1186/s40468-021-00131-8

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (Eds.). (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage Publications.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *The Journal of Marketing Theory and Practice*, *19*(2), 139–152. https://doi.org/10.2753/mtp1069-6679190202

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2018). *advanced issues in partial least squares structural equation modeling (PLS-SEM)*. Sage.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2017). *Advanced issues in partial least squares structural equation modeling*. Sage Publications.

Hasbrouck, J. E., Ihnot, C., & Rogers, G. H. (1999). "Read naturally": A strategy to increase oral reading fluency. *Reading Research and Instruction*, *39*(1), 27–37. https://doi.org/10.1080/19388079909558310

Haynes, M. (1993). Patterns and perils of guessing in second language reading. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading & vocabulary learning* (pp. 46–65). Ablex.

Haynes, M., & Baker, I. (1993). American and Chinese readers learning from lexical familiarization in English texts. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading & vocabulary learning* (pp. 130–152). Ablex.

Haziza, D. (2009). Imputation and inference in the presence of missing data. In C. R. Rao (Ed.) *Handbook of statistics* (Vol. 29, pp. 215–246). Elsevier.

Hung, H. C. M. (2017). Expertise reversal effect on reading comprehension: A case of English for specific purposes (ESP). *The Social Sciences*, *7*(1), 74–83.

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, *16*(3), 148–164. https://doi.org/10.1016/j.jslw.2007.07.005

Jalilehvand, M. (2012). The effects of text length and picture on reading comprehension of Iranian EFL students. *Asian Social Science*, *8*(3), 329–337. https://doi.org/10.5539/ass.v8n3p329

Johnson, P. (1981). Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly*, *15*(2), 169–181. https://doi.org/10.2307/3586408

Johnson, P. (1982). Effects on reading comprehension of building background knowledge. *TESOL Quarterly*, *16*(4), 503. https://doi.org/10.2307/3586468

Kameli, S., & Baki, R. (2013). The impact of vocabulary knowledge level on EFL reading comprehension. *International Journal of Applied Linguistics & English Literature*, *2*(1), 85–89. https://doi.org/10.7575/ijalel.v.2n.1p.85

Kaplan, R. B. (1966). Cultural thought patterns in intercultural education. *Language Learning*, *16*(1–2), 1–20.

Kaplan, R. B. (2005). Contrastive rhetoric. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 399–416). Routledge.

Keenan, V., Langer, P., & Medosch-Schonbeck, C. M. (1985). *Delayed retrieval following text synthesis with varied feedback* (Report No. 136). Institute of Cognitive Science, University of Colorado. https://tinyurl.com/6b94jxc2

Kelsen, B. (2016). The influence of interest and prior knowledge on EFL students' current news article/podcast reading and listening. *CAL [Magazine] Certified Akers Laboratories*, *17*(1), 80–96.

Kezhen, L. I. (2015). A study of vocabulary knowledge and reading comprehension on EFL Chinese learners. *Texas Studies in Literature and Language*, *10*(1), 33–40. https://doi.org/10.3968/n

Khataee, E., & Davoudi, M. (2018). The role of cultural schemata in inferential reading comprehension: An investigation in the Iranian EFL context. *Asean Journal of Teaching and Learning in Higher Education*, *13*(2), 11–27.

Kieras, D. E. (1978). Good and bad structure in simple paragraphs: Effects on apparent theme, reading time, and recal. *Journal of Verbal Learning and Verbal Behavior*, *17*(1), 13–28. https://doi.org/10.1016/s0022-5371(78)90496-6

Kim, K., & Clariana, R. B. (2017). Text signals influence second language expository text comprehension: Knowledge structure analysis. *Educational Technology Research and Development*, *65*(4), 909–930. https://doi.org/10.1007/s11423-016-9494-x

Kintsch, W., & Yarbrough, J. C. (1982). Role of rhetorical structure in text comprehension. *Journal of Educational Psychology*, *74*(6), 828–834. https://doi.org/10.1037/0022-0663.74.6.828

Kline, R. B. (2011). Convergence of structural equation modeling and multilevel modeling. In M. Williams & W. P. Vogt (Eds.), *The Sage handbook of innovation in social research methods* (pp. 562–589). Sage.

Kropf, M. E., & Blair, J. (2005). Eliciting survey cooperation: Incentives, self-interest, and norms of cooperation. *Evaluation Review*, *29*(6), 559–575. https://doi.org/10.1177%2F0193841X05278770

Laufer, B. (1997). The lexical plight in second language reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20–34). Cambridge University Press.

Le, V. D. (1969, June 9–14). *Some aspects of the teaching of English in Viet Nam* [Paper presentation]. Regional Seminar of the SEAMEC Regional English Language Center, Singapore. https://files.eric.ed.gov/fulltext/ED031708.pdf

Lee, R. C. (1999). *The Americas of Asian American literature*. Princeton University Press.

Lei, J. (2010). *An investigation of the effects of discourse types on Taiwanese college students reading strategy use (Publication No. 3387547)* [Doctoral dissertation, Indiana University of Pennsylvania]. Proquest Dissertations and Theses Database.

Lewis, H. (1894). *The history of the English paragraph* [Doctoral dissertation, The University of Chicago]. The University of Chicago Press.

Lin, L., Zabrucky, K., & Moore, D. (1996). The relations among interest, self-assessed comprehension, and comprehension performance in young adults. *Literacy Research and Instruction*, *36*(2), 127–139. https://doi.org/10.1080/19388079709558233

Lorch, R. F., & Chen, A. H. (1986). Effects of number signals on reading and recall. *Journal of Educational Psychology*, *78*(4), 263–270. https://doi.org/10.1037/00220663.78.4.263

Lorch, R. F., & Lorch, E. P. (1985). Topic structure representation and text recall. *Journal of Educational Psychology*, *77*(2), 137–148. https://doi.org/10.1037/0022-0663.77.2.137

5Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*(1), 111–151. https://doi.org/10.1016/0010-0285(77)90006-8

McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, *29*(2), 111–123. https://doi.org/10.1023/A:1005184709695

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading*, *12*(8), 639–646. http://www.jstor.org/stable/40011226

Mehrpour, S., & Riazi, A. (2004). The impact of text length on EFL students' reading comprehension. *Asian EFL Journal*, *6*(3), 1–13.

Meyer, B. J. F. (2003). Text coherence and readability. *Topics in Language Disorders*, *23*(3), 204–224. https://doi.org/10.1097/00011363200307000000007

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory*. North-Holland.

Meyer, B. J. F., & Freedle, R. O. (1984). Effects of discourse type on recall. *American Educational Research Journal*, *21*(1), 121–143. https://doi.org/10.3102/00028312021001121

Miccinati, J. (1975). *The effect of signal words on comprehension (Publication No. 7527484)* [Doctoral Dissertation]. ProQuest Dissertations and Theses Database.

Mikk, J. (2008). Sentence length for revealing the cognitive load reversal effect in text comprehension. *Educational Studies*, *34*(2), 119–127. https://doi.org/10.1080/03055690701811164

Mikk, J., & Kukemelk, H. (2010). The relationship of text features to the level of interest in science texts. *Trames Journal of the Humanities and Social Sciences*, *14*(1), 54–70. https://doi.org/10.3176/tr.2010.1.04

Mohammed, P. (2021). Challenges and strategies employed in comprehending short stories in English: The case of Kurdish learners. *MEXTESOL Journal*, *45*(2), 1–15.

Moon, Y. S. (2019). *Examining the effects of test characteristics on the difficulty of an EFL high-stakes reading comprehension test* [Unpublished master's thesis, Seoul University].

Namjoo, A., & Marzban, A. (2012). Text structure awareness and comprehension in EFL & ESL reading. *The Iranian EFL Journal*, *8*(6), 28–37. https://tinyurl.com/yse37nye

National Academies of Sciences, Engineering, & Medicine. (2018). *Reproducibility and Replicability in Science*. National Academies Press.

Neff, G. E. (1932). *The effect of certain changes in punctuation on speed and comprehension in reading* [Unpublished master's thesis, Ohio University].

Nelson, G. L. (1987). Culture's role in reading comprehension: A schema theoretical approach. *Journal of Reading*, *30*(5), 424–429. http://www.jstor.org/stable/40029714

Nguyen, T. T. T. (2012). *The impact of background knowledge and time constraint on reading comprehension of Vietnamese learners of English as a second language* [Unpublished Master's Thesis]. Southern Illinois University at Carbondale.

Nilagupta, S. (1977). The relationship of syntax to readability for ESL students in Thailand. *Journal of Reading*, *20*(7), 585–594. http://www.jstor.org/stable/40009837

Nitzl, C., Roldán, J. L., Carrión, G. A. C., & Hwa, C. J. (2022). *PLS2022 Prelude# 3: Mediation, moderation, and conditional mediation analysis in PLS-SEM*. https://youtu.be/YdwF-NIOYWxc?si = S0hl7RbW3Iq4Zurs

Noor, N. M. (2006). Reading academic text: Awareness and experiences among university ESL learners. *GEMA Online Journal of Language Studies*, *6*(2), 65–78.

O'Hear, M. F., Ramsey, R. N., & Baden, W. W. (1992). Measuring human interest in first-year college writing textbooks. *Literacy Research and Instruction*, *32*(1), 64–76. https://doi.org/10.1080/19388079209558106

Pathan, M., & Al-Dersi, Z. (2013). Investigating the role of short stories in overcoming the problems faced by the Libyan EFL learners in reading comprehension skill. *The Criterion An International Journal in English*, *12*, 1–8.

Pressey, L. W., & Pressey, S. L. (1921). A critical study of the concept of silent reading ability. *Journal of Educational Psychology*, *12*(1), 25–31. https://doi.org/10.1037/h0075806

Putra, B. A. W. (2012). *Cross-disciplinary effects of text factors and language of recall on reading comprehension* [Doctoral dissertation, Victoria University]. VU Research Repository. https://vuir.vu.edu.au/id/eprint/22303

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, *52*(3), 513–536. https://doi.org/10.1111/1467-9922.00193

Quan, H. (2008). The rhetorical structure approach: The role of discourse signaling cues in L2 reading comprehension. *Discourse and Intercultural Communication*, *2*, 79–95.

Raithel, S., Sarstedt, M., & Scharf, S. (2011). *On the value relevance of customer satisfaction. Multiple drivers and multiple markets*. https://doi.org/10.1007/s11747-011-0247-4

Reder, L. M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, *89*(3), 250–280. https://doi.org/10.1037/0033-295x.89.3.250

Reder, L. M., & Anderson, J. R. (1980). A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning and Verbal Behavior*, *19*(2), 121–134. https://doi.org/10.1016/s0022-5371(80)90122-x

Reder, L. M., Charney, D. H., & Morgan, K. I. (1986). The role of elaborations in learning a skill from an instructional text. *Memory & Cognition*, *14*(1), 64–78. https://doi.org/10.3758/bf03209230

Ripley, E., Macrina, F., Markowitz, M., & Gennings, C. (2010). Why do we pay? A national survey of investigators and IRB chairpersons. *Journal of Empirical Research on Human Research Ethics: JERHRE*, *5*(3), 43–56. https://doi.org/10.1525/jer.2010.5.3.43

Ritzer, A. (1994). *Investigating the knowledge of narrative and expository text structure of ESL readers*. Teachers College, Columbia University.

Roen, D. H. (1984). The effects of cohesive conjunctions, reference, response rhetorical predicates, and topic on reading rate and written free recall. *Journal of Reading Behavior*, *16*(1), 15–26. https://doi.org/10.1080/108629684 09547501

Rokni, S. J. A., & Niknaqsh, H. R. (2013). The effect of context clues on EFL learners' reading comprehension. *ELT Voices–India International Journal*, *3*(6), 54–61.

Saadatnia, M., Ketabi, S., & Tavakoli, M. (2016). EFL learners' levels of comprehension across text structures: A comparison of literal and inferential comprehension of descriptive and enumerative expository texts. *Journal of Psycholinguistic Research*, *45*(6), 1499–1513. https://doi.org/10.1007/s10936-016-9414-6

Salmani, N. M. A. (2010). The impact of formal schemata on L3 reading recall. *International Journal of Language Studies*, *4*(4), 357–372.

Salyer, M. G. (1990). *The significance of difficult vocabulary to reading in a second language* (Publication No. 9102727) [Doctoral dissertation, Michigan State University]. ProQuest Dissertations & Theses Global.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2017). Partial least squares structural equation modeling. In C. Homburg, M. Klarmann & A. Vomberg (Eds.), *Handbook of market research* (pp. 1–40). Springer.

Savage, A., & Shafiei, M. (Eds.). (2012). *Effective academic writing 1: The paragraph* (2nd ed.). Oxford University Press).

Scaglia, L. (2020). *Kant's notion of a transcendental schema: The constitution of objective cognition between epistemology and psychology*. Peter Lang.

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, *13*(1), 23–52. https://doi.org/10.1023/a:1009004801455

Shanahan, T., & Lomax, R. G. (1988). A developmental comparison of three theoretical models of the reading-writing relationship. *Research in the Teaching of English*, *22*(2), 196–212. http://www.jstor.org/stable/40171402

Sharp, A. (2002). Chinese L1 schoolchildren reading in English: The effects of rhetorical patterns. *Reading in a Foreign Language*, *14*(2), 1–22.

Shemshadsara, Z., Ahour, T., & Hadidi Tamjid, N. (2019). Raising text structure awareness: A strategy of improving EFL undergraduate students' reading comprehension ability. *Cogent Education*, *6*(1), 1–35. https://doi.org/10.1080/2331186x.2019.1644704

Shen, M. Y., & Wu, W. S. (2009). Technical university EFL learners' reading proficiency and their lexical inference performance. *Electronic Journal of Foreign Language Teaching*, *6*(2), 189–200.

Sheridan, R., Tanaka, K. M., & Hogg, N. (2019). Foreign language, local culture: How familiar contexts impact learning and engagement. *TESL-EJ*, *23*(1), 1–27.

Sherman, L. A. (1893). *Analytics of literature, a manual for the objective study of English prose and poetry*. Ginn & Company.

Shih, M. (1992). Beyond comprehension exercises in the ESL academic reading class. *TESOL Quarterly*, *26*(2), 289. https://doi.org/10.2307/3587007

Shokouhi, H., & Askari, S. (2010). The effect of guessing vocabulary in reading authentic texts among pre-university students. *Journal of Second Language Acquisition and Teaching*, *17*, 75–89.

Singer, E., & Ye, C. (2013). The use and effects of incentives in surveys. *The Annals of the American Academy of Political and Social Science*, *645*(1), 112–141. https://doi.org/10.1177%2F0002716212458082

Soper, D. S. (2022). *A-priori sample size calculator for structural equation models [software]*. https://www.danielsoper.com/statcalc

Spiro, R. J., & Taylor, B. M. (1980). *On investigating children's transition from narrative to expository discourse: The multidimensional nature of psychological text classification* (Report No. 195). Center for the Study of Reading. https://tinyurl.com/3e7fsnm2

Steffensen, M. S., Joag-Dev, C., & Anderson, R. C. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly*, *15*(1), 10–29. https://doi.org/10.2307/747429

Strube, M. J. (2000). Reliability and generalizability theory. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding MORE multivariate statistics* (pp. 23–66). American Psychological Association.

Suliman, F., Ben-Ahmeida, M., & Mahalla, S. (2019). Importance of punctuation marks for writing and reading comprehension skills. *Faculty of Arts Journal*, *13*(1), 29–53.

Summey, G. (1919). *Modern punctuation*. Oxford University Press.

Talbot, D., Ng, P., & Allan, A. (1991). Hong Kong students reading expository prose: Replication of the effects of rhetorical organization on ESL readers by Patricia Carrell. *Working Papers of the Department of English, City Polytechnic of Hong Kong*, *3*(1), 52–65.

Thaiss, C., & Zawacki, M. (2006). *Engaged writers and academic writing life*. Boynton Cook.

Thao, T. Q., & Son, T. T. (2018, June). *Factors influencing EFL reading comprehensibility among Vietnamese secondary school students: a case study* [Conference session]. Proceedings of the Language and Learning Today Conference, Vietnam. https://www.researchgate.net/profile/Tham-Duong-7/publication/332747994

Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology*, *8*(6), 323–332. https://doi.org/10.1037/h0075325

Thorndyke, P. W. (1975). *Cognitive structures in human story comprehension and memory* (ED123587) [Doctoral

dissertation, Stanford University]. https://files.eric.ed.gov/fulltext/ED123587.pdf

Van Silfhout, G., Evers-Vermeul, J., Mak, W. M., & Sanders, T. J. (2014). Connectives and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology*, *106*(4), 1036–1048. https://doi.org/10.1037/a0036293

Walters, J., & Wolf, Y. (1986). Language proficiency, text content and order effects in narrative recall. *Language Learning*, *36*(1), 47–64. https://doi.org/10.1111/j.1467-1770.1986.tb00368.x

Weaver, B. M. (2000). *Leveling books K–6: Matching readers to text*. International Reading Association.

Wold, H. (1982). Soft Modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observations: Part II* (pp. 1–54). North-Holland.

Wong, K. K. K. (2019). *Mastering partial least squares structural equation modeling (PLS-SEM) with SmartPLS in 38 hours*. IUniverse.

Xu, X., Pan, M., Dai, H., Zhang, H., & Lu, Y. (2019). How referential uncertainty is modulated by conjunctions: ERP evidence from advanced Chinese–English L2 learners and English L1 speakers. *Second Language Research*, *35*(2), 1–30. https://doi.org/10.1177%2F0267658318756948

Yali, S., & Jiliang, C. (2007). Effects of text type and test type on L2 reading comprehension test performance. *CELEA Journal*, *30*(2), 16–14.

Yorio, C. A. (1971). Some sources of reading problems for foreign-language learners[1]. *Language Learning*, *21*(1), 107–115. https://doi.org/10.1111/j.1467-1770.1971.tb00494.x

Zakaluk, B. L. (1985). *Toward a new approach to predicting text comprehensibility using inside-and outside-the-head information and a nomograph (Publication No. 8526508)* [Doctoral dissertation, University of Minnesota]. Proquest Dissertations and Theses Database.

Zakaluk, B. L., & Samuels, S. J. (1988). Toward a new approach to predicting text comprehensibility. In B. L. Zakaluk & S. J. Samuels (Eds.), *Readability: Its past, present, and future* (pp. 121–144). International Reading Association.

Zhang, X. (2008). The effects of formal schema on reading comprehension—An experiment with Chinese EFL readers. *Computational Linguistics and Chinese Language Processing*, *13*(2), 197–214.